

09/786642

PCT/JP00/04455

05.07.00

REC'D 25 AUG 2000

WIPO

PCT

日 本 国 特 許 庁

PATENT OFFICE  
JAPANESE GOVERNMENT

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日

Date of Application:

1999年 7月 7日

出 願 番 号

Application Number:

平成11年特許願第192437号

出 願 人

Applicant (s):

科学技術振興事業団

株式会社エイ・ティ・アール人間情報通信研究所

JP00/4455

18/1

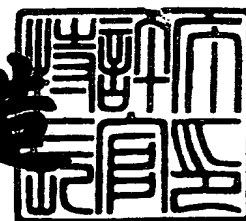
PRIORITY  
DOCUMENT

SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

2000年 8月11日

特許庁長官  
Commissioner,  
Patent Office

及川耕造



出証番号 出証特2000-3062497

【書類名】 特許願

【整理番号】 99JST24

【提出日】 平成11年 7月 7日

【あて先】 特許庁長官殿

【国際特許分類】 G10L 3/00

【発明者】

    【住所又は居所】 和歌山県海南市日方 9 3 0 番地 和歌山大学海南宿舎 1  
C - 2 2

    【氏名】 河原 英紀

【発明者】

    【住所又は居所】 京都府相楽郡精華町大字乾谷小字三平谷 5 番地 株式会  
社エイ・ティ・アール人間情報通信研究所内

    【氏名】 入野 俊夫

【特許出願人】

    【識別番号】 396020800

    【氏名又は名称】 科学技術振興事業団

    【代表者】 中村 守孝

【特許出願人】

    【識別番号】 592179296

    【氏名又は名称】 株式会社エイ・ティ・アール人間情報通信研究所

【代理人】

    【識別番号】 100089635

    【弁理士】

    【氏名又は名称】 清水 守

【手数料の表示】

    【予納台帳番号】 012128

    【納付金額】 21,000円

【提出物件の目録】

    【物件名】 明細書 1

【物件名】

図面 1

【物件名】

要約書 1

【プルーフの要否】

要

【書類名】 明細書

【発明の名称】 音源情報の抽出方法

【特許請求の範囲】

【請求項 1】 周波数から瞬時周波数への写像の不動点を用いた音源情報の抽出方法において、

各フィルタについての瞬時周波数の周波数方向の偏微分と、各フィルタ出力の周波数方向の偏微分を時間方向に偏微分した値に適切な加重をかけて、時間方向に短時間の加重付きの積分を行うことにより、各フィルタについての搬送波対雑音比の推定値を計算し、搬送波対雑音比を求め、評価量の推定値を得ることを特徴とする音源情報の抽出方法。

【請求項 2】 請求項 1 記載の音源情報の抽出方法において、前記搬送波対雑音比による評価量の推定値に基づいて、対数周波数軸上相似フィルタを基本周波数に対応する不動点の選択に用い、基本周波数についての事前情報無しに基本周波数を抽出することを特徴とする音源情報の抽出方法。

【請求項 3】 請求項 2 記載の音源情報の抽出方法において、前記対数周波数軸上相似フィルタと線形周波数軸上相似適応チャープフィルタとを組み合わせることにより基本周波数についての事前情報無しに基本周波数を抽出するとともに、該抽出された基本周波数の精度を改良することを特徴とする音源情報の抽出方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、音源情報の抽出方法に関するものである。

【0002】

【従来の技術】

瞬時周波数は時間変化信号に対する周波数概念を自然に拡大した概念である。瞬時周波数は音声のような非定常信号を表すために適した性質を多く有する。これは種々の信号処理課題に応用された；①正弦波モデルに基づいた音声符号化、②フォルマント抽出及びバンド幅推定、③有声音の調波構造の抽出、④基本周波

数（F0）の抽出、また、⑤聴覚情報処理の興味深い計算モデル等である。以下では正弦波モデルの成分正弦波の周波数、位相情報、基本周波数、およびそれらの周期性の強さ（あるいは周期成分と非周期成分の割合）等を音源情報と総称することにする。しかし、この概念に潜在する重要な可能性、特に、音声の音源情報の抽出についてはまだ十分には研究されていない。これらの側面に関する最近の研究により、瞬時周波数を利用すれば非常に優れた音源情報の抽出方法が導かれることが明らかとなった。

#### 【0003】

複数の帯域通過型フィルタの共通する通過帯域に顕著な正弦波成分がある場合、帯域通過型フィルタ出力の瞬時周波数は、中心周波数の異なったフィルタにわたって実質的には一定値をとることが知られていた。換言すれば、フィルタ中心の周波数から瞬時出力周波数までの写像は、顕著な信号周波数の近くに不動点を有する。この性質は、複合音の調波成分及び音声のフォルマントのような顕著な共振を抽出するのに用いられる。また、これらの性質と、異なった聴覚神経間の同期発火現象とが関連するであろうことが指摘されており、対応する聴覚的実体を表すものとして『音縊り（synchrony strand）』によるモデル化が行われている。しかし、これらの考えを一貫性のあるF0抽出方法としてどのようにまとめたら良いかは、明確ではなかった。

#### 【0004】

本願発明者は、近年、STRAIGHTと呼ばれる高品質の音声分析、変換、合成システムを提案している。STRAIGHTは、古典的なチャネルボコーダの概念を一般化されたピッチ同期分析に基づいて洗練したものである。ここでは、従来から用いられている述語として『ピッチ同期分析』という用語を踏襲して用いた。このように、音声情報処理の分野では、ピッチという用語が基本周波数（F0）と同じ意味で用いられている。しかし、これは、不正確な言葉の使用法である。物理的な属性を表すF0と心理的属性を表すピッチは本来別のものである。本明細書では、特に心理的属性に言及するのではない限り、『ピッチ』という用語を用いないこととする。STRAIGHT法では、F0に適応した分析が行われているため、声門の一開閉サイクルとして定義される有音声の基本周期毎

に、正確でかつ信頼性のあるF0情報が必要となる。従来から提案されている様々なF0抽出方法を適用して検討した結果、従来の方法では時間分解能についての要求条件と周波数の精度についての要求条件とを共に満たすことができないことが明らかになった。また、抽出されたF0に高速に変化する成分や不連続を含む場合には、それらの絶対値が小さくてもそのF0情報に基づいて合成された音声の知覚的品質が劣化することが分かった。さらに、知覚的に高品質な音声の合成には、無声／有声の判定が非常に大きな影響を及ぼすことが示され、数ミリ秒以内の時間的正確さが求められる場合のあることが分かった。また、逆に、特定の方向への偏りが無いのであれば、F0をゆっくりと変化させるトレンド成分には、合成された音声に対する知覚的な悪影響が無いことがわかった。

## 【0005】

## 【発明が解決しようとする課題】

現在まで、多くのF0抽出方法がある；間隔測定に基づいた時間領域アルゴリズム、スペクトルに基づいた周波数領域の方法、自己相関及びharmonic sieve（調波成分を取り出す篩い）、それらを組み合わせた方法及び生物学的に動機づけがされた方法などがある。これらの方法では、分析対象とする信号が数学的な意味での周期的信号であることを仮定している。数学的な意味での周期性に基づいて導出されたこれらの方法による推定値は、F0が時間的に一定であるような信号については、正しいF0の推定値を与える。しかし、F0が時間的に変化するような現実の音声や、複合音を構成する成分正弦波の周波数が調波性から少しだけ外れたような音を分析する場合に、従来の方法が適切なF0の推定値をあたえるかどうかは明らかではない。

## 【0006】

提案した高品質の音声変換システムでは、原音声の音源についての正確な情報に基づいて音声を変換し再合成することが必要である。したがって、この方法を改良するには、F0が時間的に変化したり調波性から外れた成分を含む信号についても合理的に適用することのできるF0抽出方法が必要となる。こうした観察が、基本成分の瞬時周波数を用いた高い時間分解能を有する正確なF0軌跡を生み出す新しいF0抽出方法の動機づけとなった。

## 【0007】

STRAIGHT法では、基本波成分を含むフィルタが最小のAM変調およびFM変調となることを仮定して瞬時周波数に基づいたF0抽出方法を導出し、用いていた。STRAIGHTで用いていたF0抽出方法は、音声と同時に記録されたEGG (Electro Glottograph) 信号を参照信号とした評価テストにおいて、妥当な性能を示した。例えば、女性話者による100文章の分析では、音声から求められたF0とEGGから求められたF0の誤差が20%以上の値を示したのは、全分析フレームの1.4%であった。また、全分析フレームの53%で、音声から求められたF0は、EGGから求められたF0の0.3%以内に入っていた。しかし、上記の最小のAM, FM変調の仮定はあいまいに定式化されており、数学的には有効でない。また、この方法では、男性の音声についてのF0の誤差の標準偏差が女性の音声の場合の2倍程度になるという問題があった。

## 【0008】

本発明は、必要な数学的基礎を提供し、上記した方法の拡張である新たなF0抽出方法を導くことである。不動点におけるフィルタ中心周波数と出力瞬時周波数との関係の偏微分についての詳細な検討は、必要な数学的基礎を提供する重要な鍵であった。これにより、瞬時周波数概念の非定常的な側面を利用する新しい一貫したF0及び音源情報抽出方法へと導かれる。

## 【0009】

本発明は、フィルタ中心周波数から出力の瞬時周波数への不動点の性質を瞬時のデータから定量的に解釈の明瞭な量として検出することができる音源情報の抽出方法を提供することを目的とする。

## 【0010】

## 【課題を解決するための手段】

〔1〕周波数から瞬時周波数への写像の不動点を用いた音源情報の抽出方法において、各フィルタについての瞬時周波数の周波数方向の偏微分と、各フィルタ出力の周波数方向の偏微分を時間方向に偏微分した値に適切な加重をかけて、時間方向に短時間の加重付きの積分を行うことにより、各フィルタについての搬送

波対雑音比の推定値を計算し、搬送波対雑音比を求め、評価量の推定値を得るようにしたものである。

【0011】

〔2〕上記〔1〕記載の音源情報の抽出方法において、前記搬送波対雑音比による評価量の推定値に基づいて、対数周波数軸上相似フィルタを基本周波数に対応する不動点の選択に用い、基本周波数についての事前情報無しに基本周波数を抽出するようにしたものである。

【0012】

〔3〕上記〔2〕記載の音源情報の抽出方法において、前記対数周波数軸上相似フィルタと線形周波数軸上相似適応チャープフィルタとを組み合わせることにより基本周波数についての事前情報無しに基本周波数を抽出するとともに、この抽出された基本周波数の精度を改良するようにしたものである。

【0013】

【発明の実施の形態】

以下、本発明の実施の形態について詳細に説明する。

【0014】

図1は本発明の実施例を示す音源情報の抽出を行うための基本周波数抽出装置のブロック図である。

【0015】

この図に示すように、入力回路1は、分析の対象となる信号 $x(t)$ を増幅、変換、分配などするために用いられる。この入力回路1において、例えば、マイクで収録された音声信号は適切なレベルに増幅された後、適切な標本化周波数でデジタル化される。デジタル化された信号は、対数周波数軸上相似フィルタ2により分析される。対数周波数軸上相似フィルタ2とは、周波数軸を対数周波数に変換してフィルタ特性を表した場合、軸上の位置のみが異なり、形状が同一のフィルタを複数、中心周波数を応用目的により定まる下限から上限まで組織的に配置したフィルタ群である。組織的な配置としては、対数周波数軸上で等間隔になるようにすることが普通である。しかし、それ以外の配置でも構わない。本発明の実験では、中心周波数を40Hzから800Hzまで、2の24乗根（約



3%の増加に相当) づつ等比的に変化させた。それぞれのフィルタは、詳細に後述する式(8)(9)(10)により求められる複素数のインパルス応答を有するフィルタである。その対数周波数軸上相似フィルタ2の出力は、瞬時周波数周波数微分回路3と不動点抽出回路6とに送られる。

#### 【0016】

瞬時周波数周波数微分回路3においては、フィルタの出力から各フィルタ出力の瞬時周波数を計算し、さらに、隣接するフィルタの出力の瞬時周波数とそれぞれのフィルタの中心周波数に基づいて、各フィルタについて瞬時周波数の周波数方向の偏微分が計算される。これは、詳細に後述する式(20)に相当する。この計算結果は、瞬時周波数時間周波数微分回路4と搬送波対雑音比計算回路5とに送られる。

#### 【0017】

瞬時周波数時間周波数微分回路4においては、瞬時周波数周波数微分回路3において求められた各フィルタについて瞬時周波数の周波数方向の偏微分の時間方向の微分を計算することにより、各フィルタ出力の瞬時周波数の周波数方向の偏微分を時間方向に偏微分した値が求められる。これは、詳細に後述する式(22)に相当する。

#### 【0018】

搬送波対雑音比計算回路5は、各フィルタについての瞬時周波数の周波数方向の偏微分と、各フィルタ出力の周波数方向の偏微分を時間方向に偏微分した値に適切な加重をかけて、時間方向に短時間の加重付きの積分を行うことにより、各フィルタについての搬送波対雑音比の推定値を計算する。それぞれの偏微分に掛ける適切な重みは、それぞれのフィルタ形状とそれぞれのフィルタの中心周波数から詳細に後述する式(12)により求められる。この加重は、分析中に変化するものではない。したがって、フィルタを設計した時点で決定することができる。こうして決定した加重の値を搬送波対雑音比計算回路5に組み込んでおけば良い。

#### 【0019】

搬送波対雑音比計算回路5の働きについては、後述する図3に具体的に例が挙

げられている。ある信号に対してその中の一つの正弦波成分を含むようなフィルタおよびその周辺のフィルタの出力から求められた量を例示している。瞬時周波数周波数微分回路 3 の出力は、図 3 の実線で表されている。瞬時周波数時間周波数微分回路 4 の出力は、図 3 の破線で示されている。これらのそれぞれを自乗し、平均して平方根を求めたものが図 3 の一点鎖線である。この一点鎖線は、瞬時周波数周波数微分回路 3 の出力と瞬時周波数時間周波数微分回路 4 の出力の全体的な傾向（振幅包絡）を表しているが、細かな振動と 1 3 5 m s 付近で 0 に非常に近くなるため実用上は使用し難い。この一点鎖線の信号を注目しているフィルタのインパルス応答の包絡により時間的に平滑化することにより、図 3 の中の点線の信号を得る。こうして求められた信号は、搬送波対雑音比の良い推定値となる。

#### 【 0 0 2 0 】

不動点抽出回路 6 は、各フィルタの中心周波数と各フィルタ出力の瞬時周波数の対応関係から不動点として安定な性質を持つものを選択し、その周波数を求める回路である。不動点の選択は、詳細に後述する式 ( 1 1 ) による。この回路自体は本発明の特徴ではない。

#### 【 0 0 2 1 】

基本周波数成分選択回路 7 は、それぞれの不動点に対応する搬送波対雑音比を比較し、最も高い搬送波対雑音比に対応する不動点を基本周波数成分として選択する。搬送波対雑音比という周波数依存性の無い客観的な尺度が推定できるようになったことにより、対数周波数軸上相似フィルタのように線形周波数軸上での形状が異なるとともに中心周波数が異なるようなフィルタ間の合理的な比較が可能になった。

#### 【 0 0 2 2 】

周期性評価回路 8 は、基本周波数成分選択回路 7 で選択された基本周波数成分の周期性の度合いを、搬送波対雑音比計算回路 5 で求められた基本周波数成分に対応する搬送波対雑音比の値に基づいて評価する回路である。ここには、3 種類の評価基準を用いることができ、それぞれ異なった 3 種類の実施例に対応する。

#### 【 0 0 2 3 】

第一の評価基準は、搬送波対雑音比をそのまま用いるものである。信号対雑音比がそのまま周期成分と非周期成分の相対的振幅を反映していると解釈するものである。

【 0 0 2 4 】

第二の評価基準は、求められた搬送波対雑音比の値をそのまま用いるのではなく、抽出された基本周波数成分の周波数の変動と振幅の変動による影響を推定して補正してから評価基準として用いる方法である。

【 0 0 2 5 】

第三の評価基準は、求められた搬送波対雑音比の値から、求められた基本周波数成分の情報に基づいて基本波だけからなる信号を作成し、その作成した信号を元の信号を分析したものと同一方法で分析して求めた作成信号の搬送波対雑音比を引いたものを、非周期成分として評価する方法である。

【 0 0 2 6 】

以上説明した部分、つまり図 1 の破線 A で囲まれた部分だけでも、高精度の音源情報分析装置として十分に利用することができる。

【 0 0 2 7 】

しかし、以下の部分、つまり図 1 の破線 B で囲まれた部分を追加することにより、さらに高精度の音源情報分析装置として利用することができる。

【 0 0 2 8 】

線形周波数軸上相似適応チャープフィルタ 9 では、後述する図 8 に示される、基本波成分選択回路により求められた基本周波数成分の基本周波数の値、周期性評価回路により求められた周期性の度合いに基づいて、周期成分が顕著な場合には、基本周波数に適応した周波数分析が行われる。ここでは、フィルタは中心周波数が線形周波数軸上で等間隔にならび、フィルタ形状も線形周波数軸上で平行移動により重なるような同じ形状を有している。このようなフィルタは、等価的に高速フーリエ変換により実現できる。また、後述する図 8 に示される、基本波成分選択回路により求められた基本周波数成分の時間微分により求められる基本周波数の瞬時周波数の変動速度に基づいて分析に先立って信号の時間軸が放物線状に変換される。この変換自体は、既に提案されている変換であるが、この変換

をこの構成の下で用いることは新しい。

#### 【0029】

瞬時周波数周波数微分回路 10 においては、フィルタの出力から各フィルタ出力の瞬時周波数を計算し、さらに、隣接するフィルタの出力の瞬時周波数とそれぞれのフィルタの中心周波数に基づいて、各フィルタについて瞬時周波数の周波数方向の偏微分が計算される。これは、詳細に後述する式 (20) に相当する。この計算結果は、瞬時周波数時間周波数微分回路 11 と搬送波対雑音比計算回路 12 とに送られる。

#### 【0030】

瞬時周波数時間周波数微分回路 11 においては、瞬時周波数周波数微分回路 10 において求められた各フィルタについて瞬時周波数の周波数方向の偏微分の時間方向の微分を計算することにより、各フィルタ出力の瞬時周波数の周波数方向の偏微分を時間方向に偏微分した値が求められる。これは、後述する式 (22) に相当する。

#### 【0031】

搬送波対雑音比計算回路 12 は、各フィルタについての瞬時周波数の周波数方向の偏微分と、各フィルタ出力の周波数方向の偏微分を時間方向に偏微分した値に適切な加重をかけて、時間方向に短時間の加重付きの積分を行うことにより、各フィルタについての搬送波対雑音比の推定値を計算する。それぞれの偏微分に掛ける適切な重みは、それぞれのフィルタ形状とそれぞれのフィルタの中心周波数から、後述する式 (12) により求められる。この加重は、分析中に変化するものではない。したがって、フィルタを設計した時点で決定することができる。こうして決定した加重の値を搬送波対雑音比計算回路 12 に組み込んでおけば良い。

#### 【0032】

不動点抽出回路 13 は、各フィルタの中心周波数と各フィルタ出力の瞬時周波数の対応関係から不動点として安定な性質を持つものを選択し、その周波数を求める回路である。不動点の選択は、後述する式 (11) による。この回路自体は、本発明の特徴ではない。

## 【0033】

帯域別周期性評価回路 14 では、それぞれのフィルタの受け持つ周波数帯域について搬送波対雑音比の値に基づいて周期性の程度を求め、それぞれの帯域の特徴を表す情報とする。

## 【0034】

基本周波数改良回路 15 では、不動点抽出回路 13 で求められた不動点の周波数の情報と、搬送波対雑音比計算回路 12 で求められた搬送波対雑音比の値を基本周波数成分選択回路 7 で求められた基本周波数の粗い推定値を参照することにより、最終的な基本周波数の推定値の平均的な誤差の期待値が最も小さくなるように統合して改良された基本周波数が求められる。

## 【0035】

なお、これらの処理と同等の処理をアナログ回路を用いて行うこともできる。その場合、入力回路 1 は、増幅ならびに分配の機能のみを有する。

## 【0036】

以下、本発明の実施例である周波数から瞬時周波数への写像の不動点及び  $F_0$  抽出方法について詳細に説明する。

## 【0037】

ここでは、フィルタ中心周波数から出力の瞬時周波数 ( $F-I F$  写像) への不動点における特徴に基づいて、信頼性のある  $F_0$  抽出方法を説明する。フィルタ包絡線のインパルス応答がガウス形の信号と 2 次のカーディナル・B スプライン (cardinal B-spline) 基底関数の畳み込みとして設定されると、不動点における  $F-I F$  写像の周波数方向の偏微分、時間周波数方向の偏微分により、顕著な正弦波上の成分 (搬送波成分) とそれ以外の成分との比率 (搬送波対雑音比) の推定値が分かる。対数周波数軸上で同じ形、また同じ間隔を有するフィルタ群を用いれば、搬送波対雑音比を基準とすることで、基本波成分を含むフィルタを選択することができる。すると信号の基本周波数は、フィルタ出力の瞬時周波数として算定される。提案した方法を音声と対応する EGG 信号とを同時に記録したデータベースを用いて評価したところ、基準となる  $F_0$  からの誤差が 20% 以上となるフレーム数は、全分析フレーム数の 1% 未満であること

が分かった。本発明により、基本周期と同程度の時間分解能での F0 軌跡の追跡が可能になる。

【0038】

以下、本発明の音源情報の抽出方法について詳細に述べる。

【0039】

〔1〕まず、このセクションでは、後のセクションで論じるために必要な概念を導入する。まず、瞬時周波数について概観する。次に、音声の駆動機構を概観した後、音声を分析するときの概念として瞬時周波数の概念が非常に優れたものであることについて述べる。

【0040】

〔1-1〕瞬時周波数

信号  $x(t)$  の瞬時周波数  $\omega(t)$  を、信号のヒルベルト変換  $H[x(t)]$  を用いて定義する。

【0041】

【数1】

$$s(t) = x(t) + jH[x(t)] \quad \dots (1)$$

【0042】

【数2】

$$\omega(t) = \frac{d \arg[s(t)]}{dt} \quad \dots (2)$$

【0043】

ここで、 $s(t)$  は解析信号であり、 $j = \sqrt{-1}$  である。この定義を直接応用するには、位相の  $2n\pi$  の不定性に伴う不連続を取り除くために位相のアンラップ操作が必要となる。位相を直接用いる必要のない方法も、こうした困難を回避するために数多く提案された。

【0044】

【数3】

$$s(t) = a(t)e^{j\phi(t)} \quad \dots (3)$$

【0045】

位相成分  $\phi(t)$  は、対応する瞬時周波数  $\omega(t)$  との以下の関係を持つ

【0046】

【数4】

$$\phi(t) = \int_{t_0}^t \omega(\tau) d\tau + \phi(t_0) \quad \dots (4)$$

【0047】

ここで、 $\phi(t_0)$  は  $t = t_0$  における初期位相である。

【0048】

瞬時周波数  $\omega(t)$  がゆっくりと変化し、信号のサンプリング間隔以下の時間内では、定数として近似することが可能であると仮定する。信号の短時間のフーリエ変換、つまり  $X(\lambda, t)$  は、以下のように定義される。

【0049】

【数5】

$$X(\lambda, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} w(t - \tau) x(\tau) e^{j\lambda\tau} d\tau \quad \dots (5)$$

【0050】

ここで、 $\omega(t)$  は時間窓を表す。各周波数地点における瞬時周波数を、2つの隣接する短時間フーリエ変換を用いて表す。

【0051】

【数6】

$$\begin{aligned} \omega(\lambda, t) &= 2f_s \arcsin \frac{|Y_d(\lambda, t)|}{2} \\ Y_d(\lambda, t) &= \frac{X(\lambda, t + \Delta t/2)}{|X(\lambda, t + \Delta t/2)|} - \frac{X(\lambda, t - \Delta t/2)}{|X(\lambda, t - \Delta t/2)|} \end{aligned} \quad \dots (6)$$

【0052】

実際は、Flanaganによる方法が計算の効率は良い。一方、上記の方程式は、離散時間信号の瞬時周波数について、概念的に簡単な解釈を提供する。この方程式において  $\omega(\lambda, t)$  を、インパルス応答  $w(t) \exp(j\lambda t)$  を有するフィルタ出力の瞬時周波数として解釈することも可能である。

〔1-2〕音声の信号モデル

有声音は、周期的構造を持つと見なされる。しかし、音声信号の基本周波数の変化は、韻律的情報を表す上で重要な役割を果たしており、高速の動きを含んでいるため厳密には周期的ではない。さらに、調波成分においてより複雑な構造が存在する。

【0053】

声門の周期的振動は、呼気流を変調して音源信号を作り出す。通常の有声音の場合には、変調された呼気流の波形には、一次導関数に周期的に不連続が生ずる。これらの不連続は、声帯の動きの開閉（時に転換点）に対応する。この不連続は高い周波数領域において高いエネルギーを有するため、こうした領域における励起の主な源となる。気流が通過することによって声帯の表面のリップルが移動するため、声門の閉止及び開き始めの時刻は、声帯の振動に完全に同期した一定の位相において生ずるとは限らない。変調された気流の波形ではエネルギーが低域に集中しているため、声門の動きは低周波領域における主な励起源である。これらの点により、調波成分の瞬時周波数は、基本周波数の正確な整数の倍数ではない。

【0054】

こうした観察によって、正弦波モデルの基本となる式として知られる以下のよう有声音のモデルが導かれる。

【0055】

【数7】

$$s(t) = \sum_{k=1}^N \sin \left( \int (k\omega_0(t) + \omega_k(t)) dt + \phi_k(0) \right) \quad \dots (7)$$

【0056】

ここで、 $\omega_0(t)$  は共通の基本周波数を表し、 $\omega_k(t)$  は  $k$  番目の成分の調波からの外れを表す。 $\phi(t)$  は初期位相を表す。

【0057】

この方程式は、基本周波数として様々な異なったものがあり得ることを示唆している。なぜなら、どの高調波成分を基準として基本周波数を計算しても構わないからである。しかし、第一の成分とより高い周波数領域の成分との間には大き



な差異が存在する。低い周波数領域における主な励起源が声帯の動きのみであるとき、高周波数領域における主な励起源は、声帯の動きとその表面上の動く波動の両方に依存する不連続の瞬間である。従って、音声信号の基本波成分を表すための基本波成分の瞬時周波数に依存することは、より簡単なモデルに対応し、かつ実際に基本的であるため、合理的なことであろう。

【0058】

## 〔2〕 F-I F 写像の不動点を用いた基本周波数推定

主要な成分以外の成分によって生じる干渉が瞬時周波数算定において主な誤差の原因であるため、正確に基本周波数を推定するために、基本波成分を分離しておくことが必要である。そうしたフィルタは、フィルタリングによる周波数及び時間方向の滲みをできるだけ避けるように設計することが必要である。

【0059】

ガウス包絡線及び2次の *cardinal B-spline* 関数の基底関数より設計されたフィルタインパルス応答により、この目的のために有用なフィルタのセットが提供される。

【0060】

## 〔2-1〕 フィルタ設計

フィルタを用いることによるスペクトルならびに時間の歪みを回避するためには、フィルタは高い時間分解能とともに隣接する調波からの干渉を十分に排除する能力を持つことが必要である。この点は、音声信号は本質的に非定常であるため、音声信号にとって不可欠である。下記のガウス包絡線より成る等方的 *Gabor* 関数は、時間周波数領域では不確定性が最小であり、時間分解能と周波数分解能の間のトレードオフの関係に関して適当な妥協点を与えるものである。“等方的”という語は、搬送波の波長及び搬送波の周波数のそれぞれについて、関数の時間周波数表示が同等の時間分解能及び周波数分解能を有するということを表している。

【0061】

【数 8】

$$w(t) = \frac{1}{\tau_0} e^{-\pi(t/\tau_0)^2} \quad \dots (8)$$

【0 0 6 2】

【数 9】

$$W(\omega) = \frac{\tau_0}{\sqrt{2\pi}} e^{-\pi(\omega/\omega_0)^2} \quad \dots (9)$$

【0 0 6 3】

ここで、 $W(\omega)$  はインパルス応答  $w(t)$  のフーリエ変換、また  $\omega_0 = 2\pi f_0$  はフィルタの中心周波数である。

【0 0 6 4】

2 次の cardinal B-spline 関数の基底関数を等方的ガウス包絡線関数で畳み込むことにより、隣接する調波成分によって生ずる干渉を抑制するために隣接調波の周波数の付近に 2 次の零点が加えられる。

【0 0 6 5】

【数 1 0】

$$w_p(t) = e^{-\pi\left(\frac{t}{\tau_0}\right)^2} * h(t/\tau_0)$$

$$h(t) = \begin{cases} 1 - |t| & |t| < 1 \\ 0 & \text{otherwise} \end{cases} \quad \dots (10)$$

【0 0 6 6】

ここで \* は畳み込みを表す。

【0 0 6 7】

〔2-2〕 正弦波状の成分の抽出

支配的な正弦波状信号のみがフィルタの実効的な通過域の中にあると仮定しよう。この時、フィルタ出力の瞬時周波数は周波数、つまり支配的正弦波状の成分の  $\omega_d$  によって決定される。換言すれば、フィルタ出力の瞬時周波数は、そうしたフィルタが共通の支配的正弦波状の成分を共有するとき、ほとんど同一である。正弦波状の成分の周波数を  $\omega_s(t)$  で表す。これによって、 $\omega_s(t)$  の近傍に不動点が存在するようになる。 $\omega_s(t)$  より低い中心周波数を持つフィル

タの出力の瞬時周波数は、その中心周波数よりも高い。その一方、 $\omega_s(t)$  より高い中心周波数を持つフィルタの出力の瞬時周波数は、その中心周波数よりも低い。中心周波数がこれら2つの中心周波数の間で変化するときに出る瞬時周波数は連続的に変化するため、フィルタ出力の瞬時周波数がその中心周波数と一致する地点が存在するが、これが不動点である。不動点の上側のフィルタと下側のフィルタの中心周波数の不動点の周波数からのずれは任意に小さくできるので、結局、不動点の周波数は $\omega_s(t)$ と一致する。

【0068】

フィルタの中心周波数を $\lambda$ で表し、また $\omega_i(\lambda, t)$ でフィルタ出力の瞬時周波数を表す。このようにすると、次の式で定義された不動点の集合は、信号に含まれる正弦波状の成分の候補を与える。

【0069】

【数11】

$$\Lambda(t) = \{ \lambda \mid \omega_i(\lambda, t) = \lambda, \omega_i(\lambda - \varepsilon, t) - (\lambda_n - \varepsilon) > \omega_i(\lambda + \varepsilon, t) - (\lambda_n + \varepsilon) \} \quad \dots (11)$$

【0070】

ここで、 $\varepsilon$ は任意の小さな定数を表す。

【0071】

〔3-3〕搬送波対雑音比の推定

支配的な正弦波状の成分のみが実効的な通過域に存在するとき、出力瞬時周波数は、正弦波状の成分の周波数と全く同じである。背景雑音が支配的な正弦波状の成分に関して十分に小さい場合、不動点の近くのフィルタ出力の瞬時周波数の誤差は、正弦波状の成分として表された背景雑音の加重和によって近似される。この雑音成分が不動点のまわりのフィルタの実効的な通過域において均一に分布していると仮定するならば、支配的な正弦波状の成分の周波数とフィルタ出力の瞬時周波数の誤差の分散は、背景雑音の相対的な誤差の分散に比例する。なお、相対的な誤差の分散を平均自乗誤差として表したものの逆数が搬送波対雑音比である。背景雑音の相対的な誤差の分散は、以下の式を用いて、不動点におけるF-I F写像の周波数偏微分及び時間周波数偏微分から推定することができる。

【0072】

相対的誤差分散を  $\sigma^2$  で表す。

【0073】

【数12】

$$\begin{aligned}\bar{\sigma}^2 &= c_a \left( \frac{\partial \omega_i(t, \lambda)}{\partial \lambda} \right)^2 + c_b \left( \frac{\partial^2 \omega_i(t, \lambda)}{\partial t \partial \lambda} \right)^2 \\ c_a &= \frac{1}{\int \left( \delta \frac{dW_p(\omega)}{d\omega} \Big|_{\omega=\delta} \right)^2 d\delta} \\ c_b &= \frac{1}{\int \left( \delta^2 \frac{dW_p(\omega)}{d\omega} \Big|_{\omega=\delta} \right)^2 d\delta} \quad \dots (12)\end{aligned}$$

【0074】

ここで、 $W_p(\omega)$  はフィルタ応答  $\omega_p(t)$  のフーリエ変換を表す。実際には、相対的誤差分散の確かな推定値を求めるためには時間的平滑化を取り入れることが必要である。

【0075】

〔2-4〕基本波成分の選択

システムが時間分解能と周波数分解能の間の最良の妥協点を実現するためには、目的とする主要な正弦波状の成分に関する情報を用いてフィルタを設計することが必要である。基本周波数抽出のそれを設計するには、基本周波数についての事前の情報がまた必要とされる。しかし、そのような情報は、分析するために予め利用することはできない。こうした困難を回避する一つの方法として、体系的に設計された形状と中心周波数を持つ一連のフィルタを用いる方法がある。

【0076】

一連のフィルタが対数周波数軸上に等しい周波数間隔を持ち、対数周波数軸上に同じ形状を持つと仮定する。フィルタの間隔が十分に密であれば、事実上、あらゆる不動点はフィルタ中心に位置する。すると、基本周波数に対応する不動点から構成されるフィルタが最小の相対的誤差分散を有する。これは、他のフィルタが実効的な通過域の中に必然的に複数の調波成分や雑音成分を含んでしまうからである。換言すれば、相対的誤差分散が最小であることは、不動点が基本波成

分を表すという証拠である。この議論の進め方は、本願発明者が以前の発明で用いた『基本波らしさ』の概念を導いた時と同様である。しかし、以前の考えは、FMとAMの大きさの合計を測定する、直観的に取り入れた方法に基づいたものであり、確実な数学的基礎に基づいたものではない。また、相対的誤差分散は、周波数の推定誤差と直接的に対応していることもあって、より適切である。

【0077】

以上の検討に基づき、F0の事前情報に依存しない基本波成分の選択手続きは、以下のようにまとめられる。

【0078】

●ステップ1：対数関数軸上に等間隔に置かれた中心周波数を持つ一連のフィルタを用意する。中心周波数はF0の存在しうる範囲を覆わなければならない（すなわち40Hz～800Hz）。間隔は十分に密でなくてはならない（すなわち1オクターブにつき24フィルタ）。

【0079】

●ステップ2：分析対象の信号を用意したフィルタへ送り込む。

【0080】

●ステップ3：各フィルタ出力につき瞬時周波数を算定する。

【0081】

●ステップ4：選択基準を用いて不動点を抽出する〔式（11）〕。

【0082】

●ステップ5：各不動点につき相対的誤差分散を算定する〔式（12）〕。

【0083】

●ステップ6：各分析フレームにおいて、最小の相対的誤差分散を有する不動点を選択する。こうして選択された不動点は、基本波成分の最も有力な候補である。

【0084】

基本周波数は、抽出した基本波成分の瞬時周波数として推定される。

【0085】

実際には、基本波成分を選択する最終ステップは、録音の時の環境騒音等の影

響を防ぐために挿入される高域通過フィルタの影響や低い周波数における信号対雑音比の劣化の影響によって、基本波成分に対応する相対的誤差分散の大きさが十分に小さくならないため、失敗することがある。この問題の影響は、相対的誤差分散が十分に小さな部分から求められる F 0 軌跡を、その前後に連続性を追跡しながら探索して延長することによって軽減することができる。

【0086】

〔2-5〕余分の正弦波状の成分によって生じた干渉

顕著な正弦波状の成分の一つに中心を置くフィルタ出力信号を、下記の方程式によって近似することができる。  $\varepsilon \ll 1$  と仮定する。

【0087】

【数13】

$$s(t) = g(\omega - \omega_h) e^{j\omega_h t} + \varepsilon g(\omega - \omega_h + \delta) e^{j(\omega_h + \delta)t} \quad \dots (13)$$

【0088】

【数14】

$$= e^{j\omega_h t} g(\omega - \omega_h) \left( 1 + \frac{\varepsilon g(\omega - \omega_h + \delta)}{g(\omega - \omega_h)} e^{j\delta t} \right) \quad \dots (14)$$

【0089】

$g(\omega)$  は、 $\omega = 1$  において最大値 1 を持つとする。周波数領域の重み関数  $g(\omega)$  は滑らかな連続関数であり、 $\omega = 0$  の周辺には特異点がないものとする。この時、0 付近の  $g(\omega)$  の Taylor 展開は、 $\omega \ll 1$  なら、 $g(\omega) \doteq 1$  であることが分かる。これらの仮定を用いれば、上記 (14) 式は次のように近似される。

【0090】

【数15】

$$s(t) \simeq e^{j\omega_h t} (1 + \varepsilon g(\omega - \omega_h + \delta) e^{j\delta t}) \quad \dots (15)$$

【0091】

ここで、瞬時周波数を調べるためには、この方程式を極形式に書き換える必要がある。

【0 0 9 2】

【数 1 6】

$$\begin{aligned} s(t) &\simeq e^{j\omega_h t} (1 + \varepsilon g(\omega - \omega_h + \delta) e^{j\delta t}) \\ &= \sqrt{1 + 2\varepsilon g(\omega - \omega_h + \delta) \cos \delta t + \varepsilon^2 g^2(\omega - \omega_h + \delta)} e^{j \tan^{-1} \frac{\varepsilon g(\omega - \omega_h + \delta) \sin \delta t}{1 + \varepsilon g(\omega - \omega_h + \delta) \cos \delta t}} e^{j\omega_h t} \end{aligned} \quad \dots (16)$$

【0 0 9 3】

条件を  $\omega \ll 1$  及び  $\varepsilon \ll 1$  と仮定するので、その方程式はさらに近似される。

【0 0 9 4】

【数 1 7】

$$\begin{aligned} s(t) &\simeq (1 + \varepsilon g(\omega - \omega_h + \delta) \cos \delta t) e^{j \tan^{-1} \varepsilon g(\omega - \omega_h + \delta) \sin \delta t} e^{j\omega_h t} \\ &\simeq (1 + \varepsilon g(\omega - \omega_h + \delta) \cos \delta t) e^{j \varepsilon g(\omega - \omega_h + \delta) \sin \delta t} e^{j\omega_h t} \\ &= (1 + \varepsilon g(\omega - \omega_h + \delta) \cos \delta t) e^{j\omega_h t + j \varepsilon g(\omega - \omega_h + \delta) \sin \delta t} \end{aligned} \quad \dots (17)$$

【0 0 9 5】

信号  $s(t)$  の位相関数  $\phi(t)$  を下記のように近似する。

【0 0 9 6】

【数 1 8】

$$\phi(t) \simeq \omega_h t + \varepsilon g(\omega - \omega_h + \delta) \sin \delta t \quad \dots (18)$$

【0 0 9 7】

これは、干渉信号により位相変調が生ずることを示している。

【0 0 9 8】

信号  $s(t)$  の瞬時周波数  $\omega_i(t)$  は位相関数の時間導関数より導かれる。  
それは下記のようなになる。

【0 0 9 9】

【数 19】

$$\begin{aligned}\omega_i(t) &= \frac{d\phi(t)}{dt} \\ &\approx \frac{d}{dt}(\omega_h t + \varepsilon g(\omega - \omega_h + \delta) \sin \delta t) \\ &= \omega_h(t) + t \frac{d\omega_h(t)}{dt} + \varepsilon \delta g(\omega - \omega_h + \delta) \cos \delta t\end{aligned}\quad \dots (19)$$

【0100】

〔2-6〕搬送波対雑音比の実際的な推定方法

ここで求めたいのは、問題の正弦波状の成分に対する搬送波対雑音比である。それを瞬時値だけに基づいて計算できることが望ましい。換言すれば、特定の帯域通過フィルタの通過域内での  $\varepsilon$  の平均を求めるのである。つまり、基本的な考えは、 $\sin^2 + \cos^2 = 1$  の関係を用いて、 $\omega_i(t)$  における正弦波状の変動を除去する方法を導くことである。不動点における幾何学的属性は、これを達成する鍵になる。

【0101】

〔2-6-1〕周波数偏微分

瞬時周波数  $\omega_i(t)$  の周波数に関する偏微分から次式が得られる。

【0102】

【数 20】

$$\begin{aligned}\frac{\partial \omega_i(t, \omega)}{\partial \omega} &\simeq \frac{\partial}{\partial \omega} \left( \omega_h(t) + t \frac{d\omega_h(t)}{dt} + \varepsilon \delta g(\omega - \omega_h + \delta) \cos \delta t \right) \\ &= \left( \frac{\partial g(\omega - \omega_h + \delta)}{\partial \omega} \right) \varepsilon \delta \cos \delta t \\ &= \left( \frac{dg(\omega - \omega_h + \delta)}{d\omega} \right) \varepsilon \delta \cos \delta t \\ &= \left. \frac{dg(\omega)}{d\omega} \right|_{\omega=\delta} \varepsilon \delta \cos \delta t\end{aligned}\quad \dots (20)$$

【0103】

干渉する成分が一つだけの場合、 $t_0 = 2\pi / \delta$  で決まる一周期分を観測するだけで  $\varepsilon$  の値を推定することが可能である。しかし、一般に、同時に複数の成分が存在し得る。



【0104】

〔2-6-2〕時間周波数偏微分

時間に関する偏微分を求めることにより、前の余弦波位相を有する信号に対応する正弦波位相の対応物を求めるのは、もっともと思われる。

【0105】

【数21】

$$\begin{aligned}\frac{\partial \omega_i(t, \omega)}{\partial t} &\simeq \frac{\partial}{\partial t} \left( \omega_h(t) + t \frac{d\omega_h(t)}{dt} + \varepsilon \delta g(\omega - \omega_h + \delta) \cos \delta t \right) \\ &= \frac{d\omega_h(t)}{dt} + \frac{d\omega_h(t)}{dt} + t \frac{d^2 \omega_h(t)}{dt^2} - \varepsilon \delta^2 g(\omega - \omega_h + \delta) \sin \delta t \\ &= 2 \frac{d\omega_h(t)}{dt} + t \frac{d^2 \omega_h(t)}{dt^2} - \varepsilon \delta^2 g(\omega - \omega_h + \delta) \sin \delta t \quad \dots (21)\end{aligned}$$

【0106】

求める正弦波位相変数を第3項で得る。しかし、音声のような信号は基本周波数が高速に変化し、その変化についての事前情報も得られないため、初めの2項を除去することはできない。

【0107】

次の段階は、方程式(21)の周波数に関する偏微分を導くことである。これは、以下ようになる。

【0108】

【数22】

$$\begin{aligned}\frac{\partial^2 \omega_i(t, \omega)}{\partial t \partial \omega} &\simeq \frac{\partial}{\partial \omega} \left( 2 \frac{d\omega_h(t)}{dt} + t \frac{d^2 \omega_h(t)}{dt^2} - \varepsilon \delta^2 g(\omega - \omega_h + \delta) \sin \delta t \right) \\ &= - \left( \frac{dg(\omega - \omega_h + \delta)}{d\omega} \right) \varepsilon \delta^2 \sin \delta t \\ &= - \frac{dg(\omega)}{d\omega} \Big|_{\omega=\delta} \varepsilon \delta^2 \sin \delta t \quad \dots (22)\end{aligned}$$

【0109】

これは、正弦波位相で変化する成分のみから成る。

【0110】

〔3〕以下、具体的数値例について述べる。

## 【0111】

人工的信号及び実際の音声試料を用いた分析例について述べる。

## 【0112】

## 〔3-1〕付加的な白色雑音を持つインパルス列

図2にフィルタ中心周波数から出力瞬時周波数への写像を示す。200 Hz パルス列と白色雑音（S/Nは20 dB）との合成信号を、対数周波数軸上で等間隔に配置したフィルタを用いて分析する。200 Hzに対応する不動点付近の瞬時周波数は均一のままである点に注意されたい。他の不動点は、このような安定性を示していない。

## 【0113】

図3に搬送波対雑音比の計算に用いられる様々な中間的な変数の値と最終的に得られた結果の例を示す。この図において、それらの平方根の値を図3上に記入する。実線で示される周波数偏微分の系列と破線の時間周波数偏微分との間に $\pi/2$ の位相差がうまく導入されていることに注意されたい。また、135 ms 付近の点において、周波数偏微分と時間周波数偏微分の重み付き自乗平均値に、成分正弦波の間の干渉に起因する鋭い窪みが生じていることが分かる。この重み付き自乗平均値に前述の平滑化を適用することにより、滑らかな搬送波対雑音比の推定値が求められる。

## 【0114】

図4に搬送波対雑音比の時間一周波数（時間-チャンネル番号）表示を画像として表示する。また、図4では求められた不動点をその上に重ねて表示している。図では暗さが搬送波対雑音比の大きさに対応しており、暗いほど搬送波対雑音比が大きい。

## 【0115】

200 Hz 付近の抽出した不動点のほぼ全ては、基本波成分に対応する。他の不動点の中には200 Hz 付近に位置するものはない。100 Hz 未満の領域では抽出した不動点が無作為に分布しているが、それが互いに近づく傾向は弱い。より周波数の高い領域では、不動点は調波周波数付近にとどまる傾向がある。

## 【0116】

図5に瞬時周波数と搬送波対雑音比によって張られる平面上での不動点の分布を示す。基本成分に対応する不動点は、明らかに別個のものである。調波周波数付近の不動点の搬送波対雑音比は、調波周波数において最大値を示すという点に注意されたい。このような現象が生ずるのは、隣接する調波成分が同程度の大きさで混合される場合に相互の干渉が非常に大きくなるからである。

【0117】

図6に最小点と残りの点の搬送波対雑音比の分布を示す。基本波成分に対応する不動点は、はっきりと区別できる分布を有することが分かる。

【0118】

### 〔3-2〕持続母音

図7に男性の話者による持続した日本語の母音／a／を入力信号とした場合の、中心周波数から瞬時周波数への写像を示す。話者には、持続母音の発声に際しては一定（約130Hz）の基本周波数を保つよう指示した。信号の標本化周波数は22050Hz、量子化ビット数は16bitであった。パルス列の場合のように、基本周波数に対応する不動点付近では、写像は実質的には平坦である。

【0119】

図8に瞬時周波数と搬送波対雑音比によって張られる平面上での不動点の分布を示す。基本波成分に対応する不動点は、130Hz付近に位置する。

【0120】

図9に瞬時周波数と搬送波対雑音比の散布図を示す。この図から基本波成分付近の不動点が非常に小さな搬送波対雑音比を有することは明らかである。パルス列の場合のように、調波成分付近の不動点は、調波周波数において最大の搬送波対雑音比を示す。基本波成分についての搬送波対雑音比は約40dBであり、持続母音のF0が非常に安定していることを示す。

【0121】

図10に度数分布表示における同じデータを示す。この図から分布が分離していることは明らかである。

【0122】

### 〔3-3〕自然な韻律を有する母音連鎖

図11に男性話者による、連続的に発音された母音連鎖より抽出した不動点の時間周波数散布図を示す。以前の結果と同様、この図に基本波成分に対応する軌跡が滑らかに連続する不動点の集まりとしてはっきり見える。第一のフォルマントに対応する不動点は、500msから700msのあたりにはっきりと見える。

図12に不動点の搬送波対雑音比の時間経過を示す。この図では、有声音の部分が明瞭に分かる。有声音部分では、基本波成分のみが十分に大きな搬送波対雑音比を示している。

【0123】

図13に瞬時周波数と搬送波対雑音比の分布を示す。この図13と図11とを併せて考えれば、先読み用のバッファを用いることで、信頼性の高いF0追跡アルゴリズムを容易に実現することができる。

【0124】

〔3-4〕 同時EGG記録を用いたセンテンス（文）データベース

図14に基本周波数推定における誤差分布を示す。図の横軸は、音声信号から求められたF0とEGG信号から求められたF0の周波数の比を百分率で表したものである。横軸上の100%の位置は、誤差が0である場合に対応している。図14（a）は男性話者による基本周波数推定における誤差を、図14（b）は女性話者による基本周波数推定における誤差をそれぞれ示している。これらの図によると、男性話者の誤差は、女性話者のそれよりも大きいことが分かる。

【0125】

【表 1】

	男 フレーム数 (全フレームに対する割合：%)	
総フレーム数	1 5 6 1 0 2	
20%以上の誤差	7 1 2	(0. 4 5 6 1%)
5%以上の誤差	1 0 9 6 3	(7. 0 2 3%)
1%以上の誤差	6 4 9 2 6	(41. 5 9%)
ハーフピッチ誤差	6 3	(0. 0 4 0 3 6%)
ダブルピッチ誤差	2 8 1	(0. 1 8%)
	女 フレーム数 (全フレームに対する割合：%)	
総フレーム数	2 4 9 6 4 1	
20%以上の誤差	1 8 1	(0. 0 7 2 5%)
5%以上の誤差	2 5 7 7	(1. 0 3 2%)
1%以上の誤差	2 6 1 1 1	(10. 4 6%)
ハーフピッチ誤差	4 6	(0. 0 1 8 4 3%)
ダブルピッチ誤差	1 8	(0. 0 0 7 2 1%)

## 【0126】

表1に基本周波数抽出における誤差の統計を示す。結果の中にはEGG信号の分析の誤差も含まれていることに注意する必要があるものの、これは非常に良い結果である。この結果は、不動点に基づいたF0推定法において基本波成分のみを用いた場合の性能の上限であるとみなすことができる。女性のデータはほぼ満足の行くものであるとの結論が出せるが、男性のデータはさらに改善が必要であると言える。図1の破線Bの部分は、このような場合の推定結果を改良するために用いられる。

## 【0127】

なお、本発明は上記実施例に限定されるものではなく、本発明の趣旨に基づいて種々の変形が可能であり、これらを本発明の範囲から排除するものではない。

## 【0128】

【発明の効果】

以上、詳細に説明したように、本発明によれば、以下のような効果を奏することができる。

【0 1 2 9】

(A) 信号の中にある正弦波成分を精度よく確実に抽出することができるとともに、抽出された成分の影響を短時間の観測値から定量的に求めることができる。

【0 1 3 0】

(B) 分析合成音声を合成するための高品質の音源情報（基本周波数及び周期性についての情報）の抽出を行うことができる。

【0 1 3 1】

(C) 楽器音などの周期性を有する音の分析においても、周期性の確からしさを客観的な指標として求めることができたため、分析結果に基づいて楽器音を変換・再合成する場合の、高品質の音源情報として用いることができる。また、汎用の分析装置として一般の信号の周期性の分析にも用いることができる。

【0 1 3 2】

(D) 定量的に解釈の明確な量が求められるため、対数周波数軸上相似フィルタと線形周波数軸上相似適応チャープフィルタのように構造の異なったフィルタによる結果を有効に統合することができる。

【0 1 3 3】

(E) 搬送波対雑音比の推定値は、そのまま、帯域フィルタあるいは周波数分析結果の評価に用いることができる。

【図面の簡単な説明】

【図 1】

本発明の実施例を示す音源情報の抽出を行うための基本周波数抽出装置のブロック図である。

【図 2】

本発明の実施例を示すフィルタ中心周波数から出力瞬時周波数への写像を示す図である。

【図 3】

本発明の実施例を示す搬送波対雑音比を計算する過程の中間結果及び最終結果を示す図である。

【図 4】

本発明の実施例を示す時間一チャネル平面における搬送波対雑音比と不動点の分布を示す図である。

【図 5】

本発明の実施例を示すフィルタ出力の瞬時周波数と搬送波対雑音比の分布を示す図である。

【図 6】

本発明の実施例を示す搬送波対雑音比の度数分布を示す図である。

【図 7】

本発明の実施例を示すフィルタの中心周波数から出力の瞬時周波数への写像を示す図である。

【図 8】

本発明の実施例を示す時間一チャネル平面における搬送波対雑音比と不動点の分布を示す図である。

【図 9】

本発明の実施例を示すフィルタ出力の瞬時周波数と搬送波対雑音比の分布を示す図である。

【図 1 0】

本発明の実施例を示す搬送波対雑音比の度数分布を示す図である。

【図 1 1】

本発明の実施例を示す時間一チャネル平面における搬送波対雑音比と不動点の分布を示す図である。

【図 1 2】

本発明の実施例を示す搬送波に対する相対的雑音振幅の時間的分布を示す図である。

【図 1 3】

本発明の実施例を示すフィルタ出力の瞬時周波数と搬送波対雑音比の分布を示す図である。

【図 1 4】

本発明の実施例を示す F 0 推定誤差の分布を示す図である。

【符号の説明】

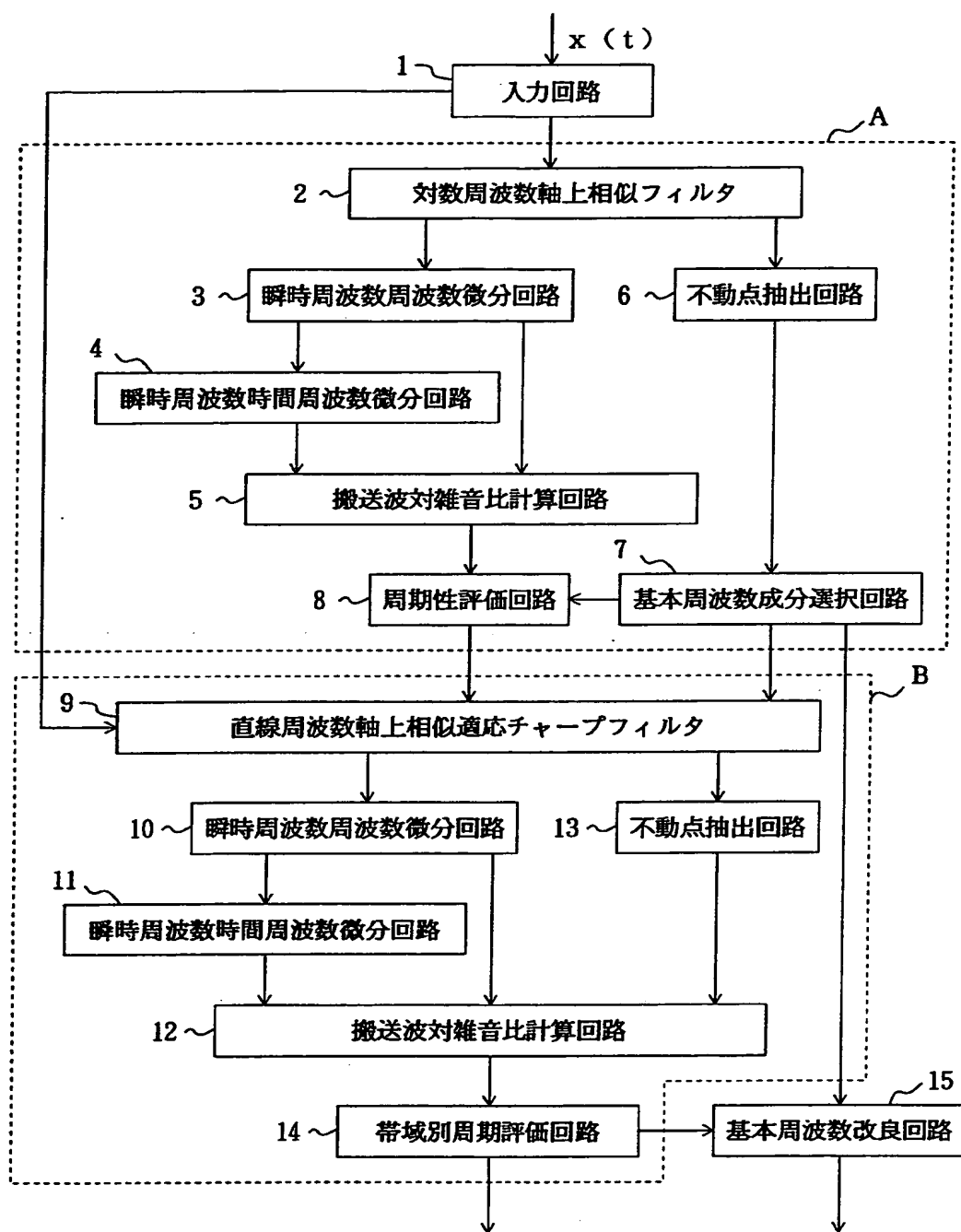
- 1     入力回路
- 2     対数周波数軸上相似フィルタ
- 3, 1 0     瞬時周波数周波数微分回路
- 4, 1 1     瞬時周波数時間周波数微分回路
- 5, 1 2     搬送波対雑音比計算回路
- 6, 1 3     不動点抽出回路
- 7     基本周波数成分選択回路
- 8     周期性評価回路
- 9     線形周波数軸上相似適応チャープフィルタ
- 1 4     帯域別周期性評価回路
- 1 5     基本周波数改良回路



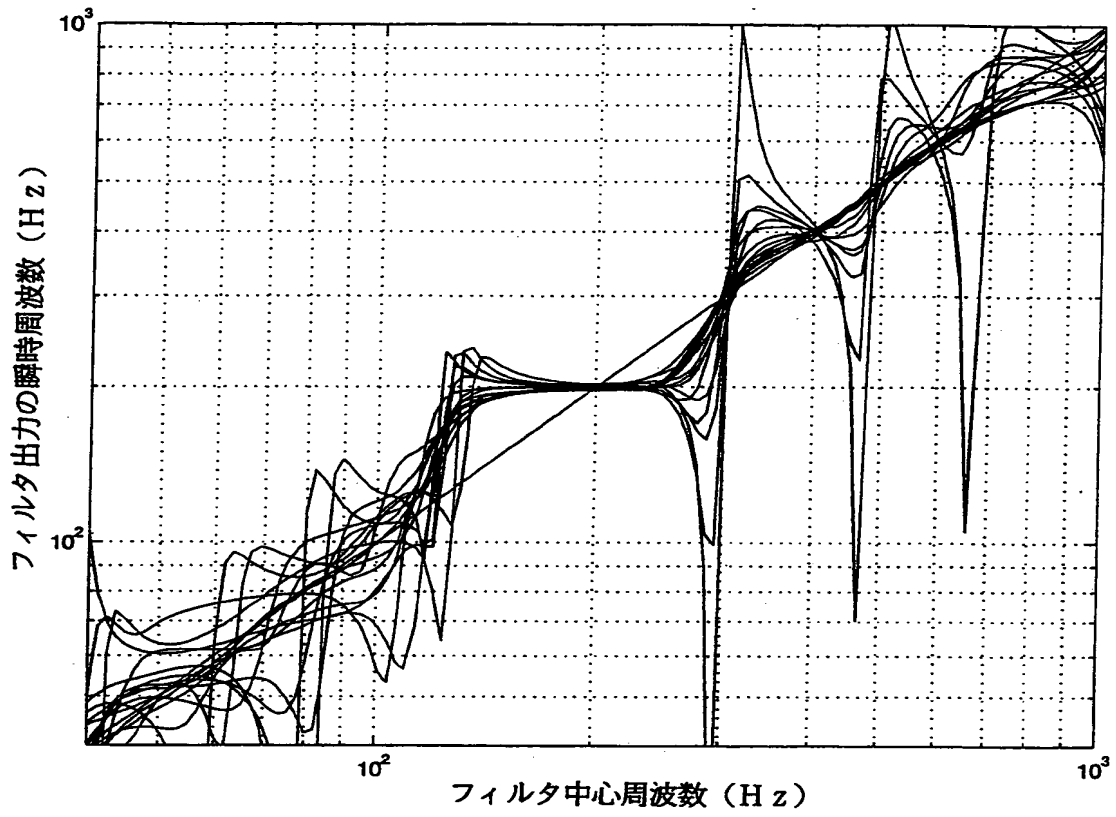
【書類名】

図面

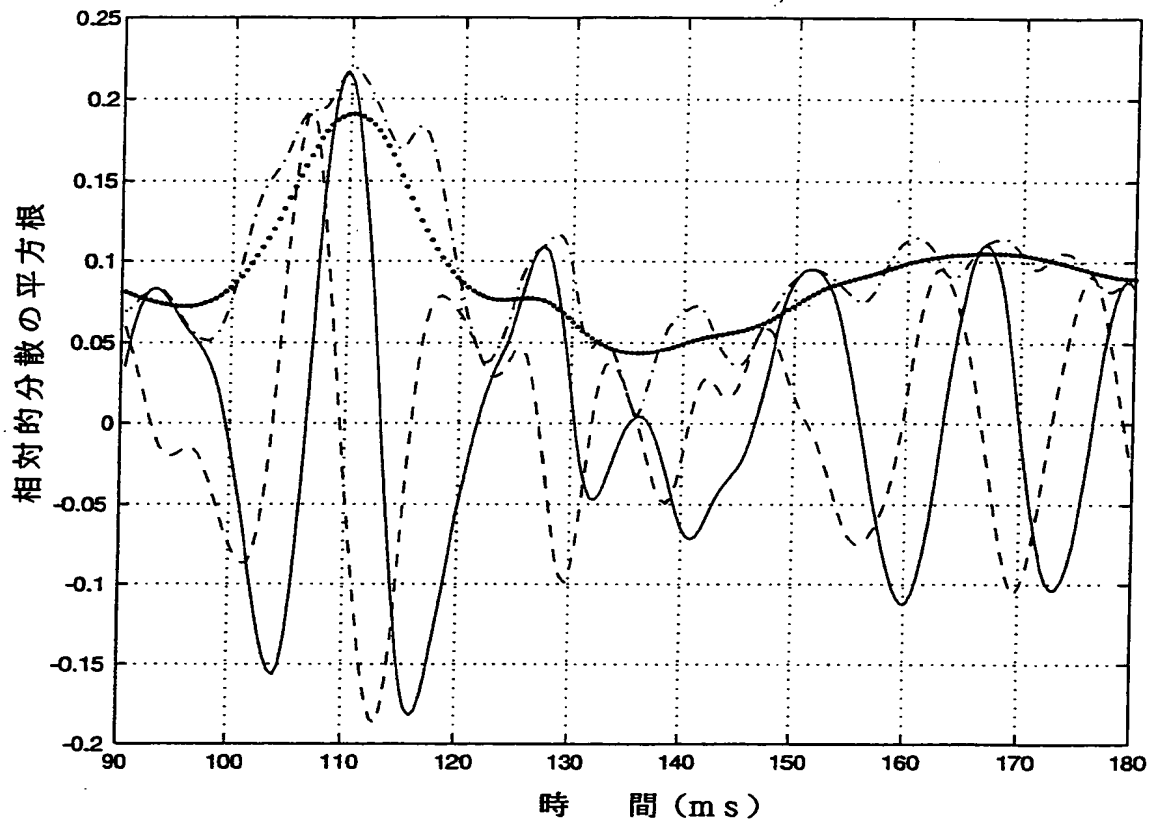
【図 1】



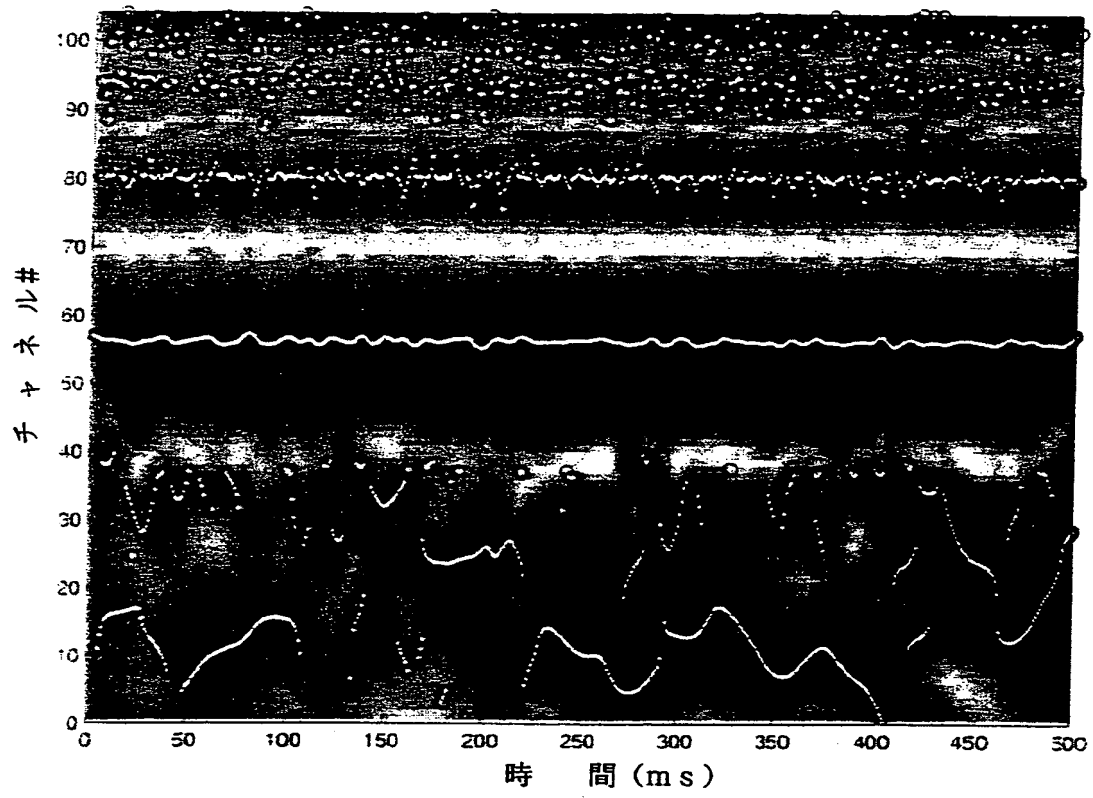
【図 2】



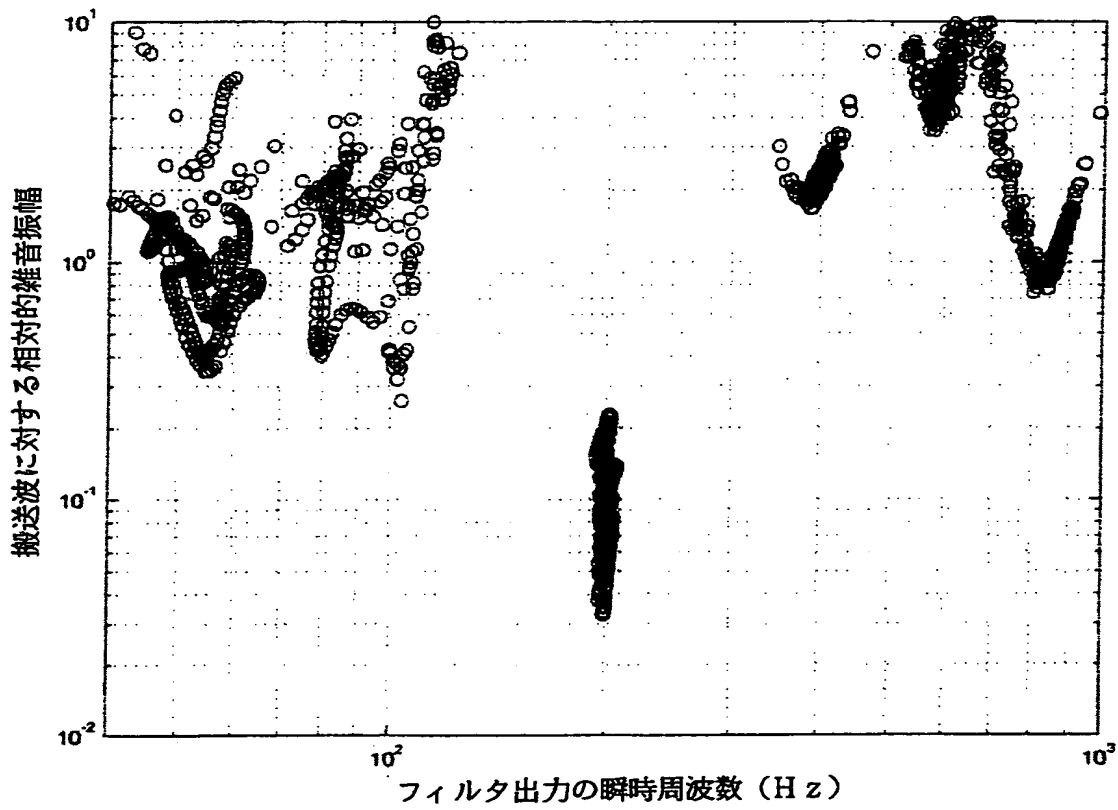
【図 3】



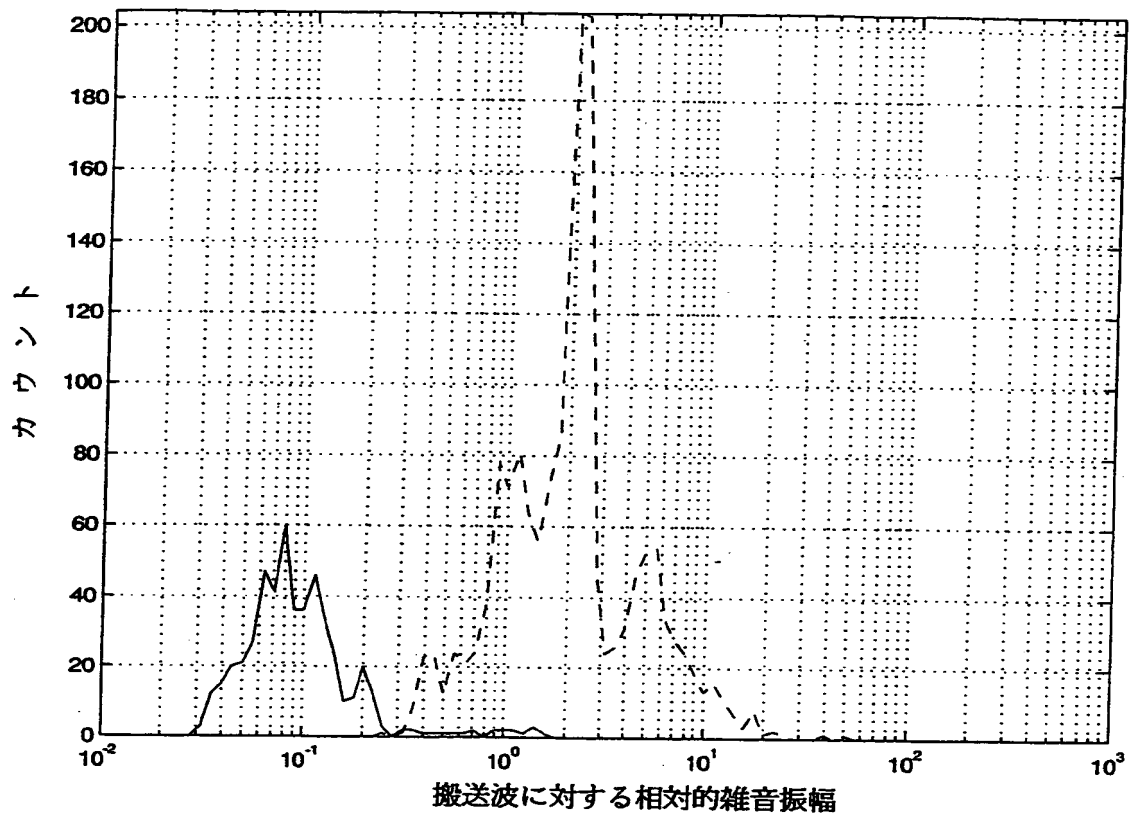
【図4】



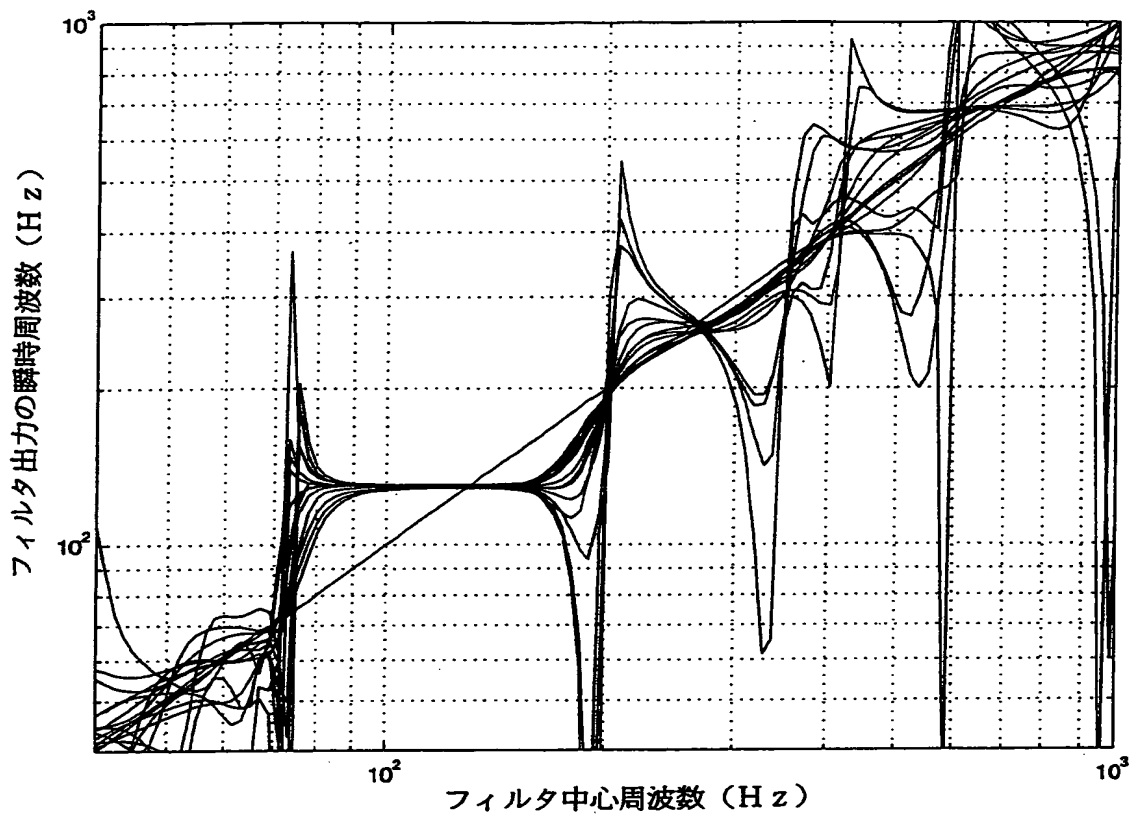
【図 5】



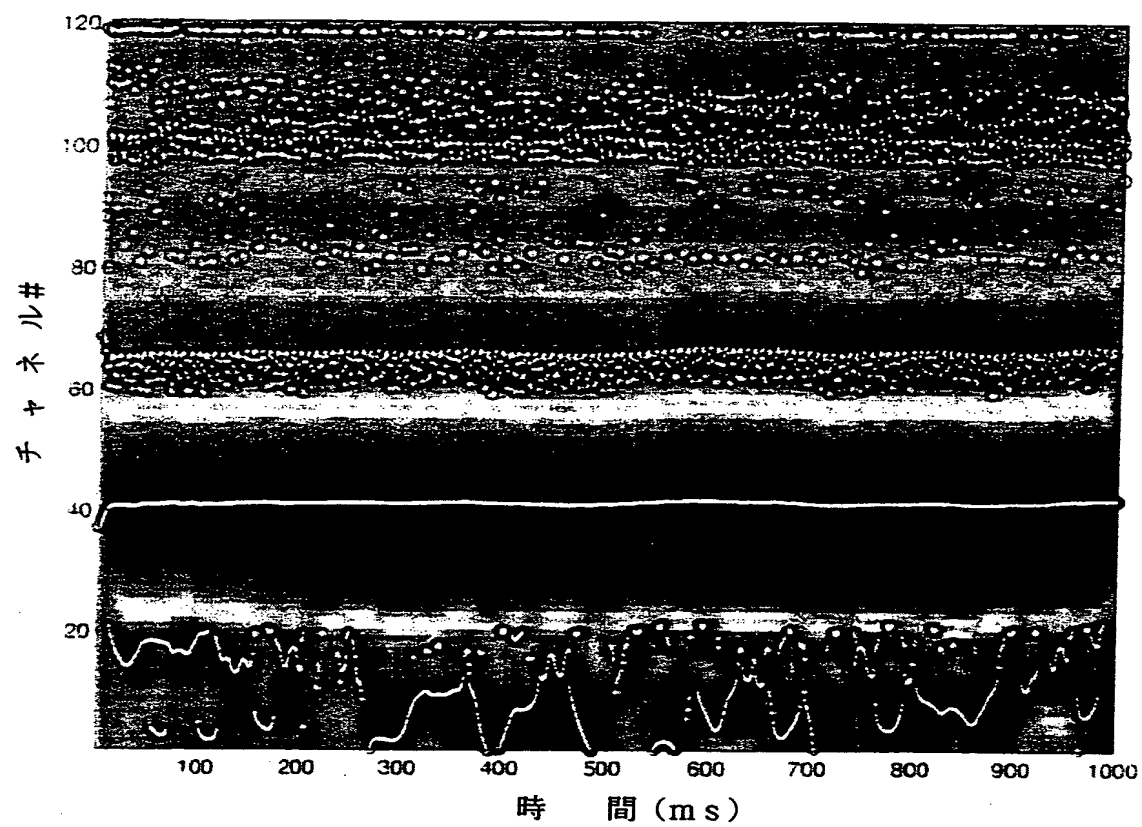
【図 6】



【図 7】

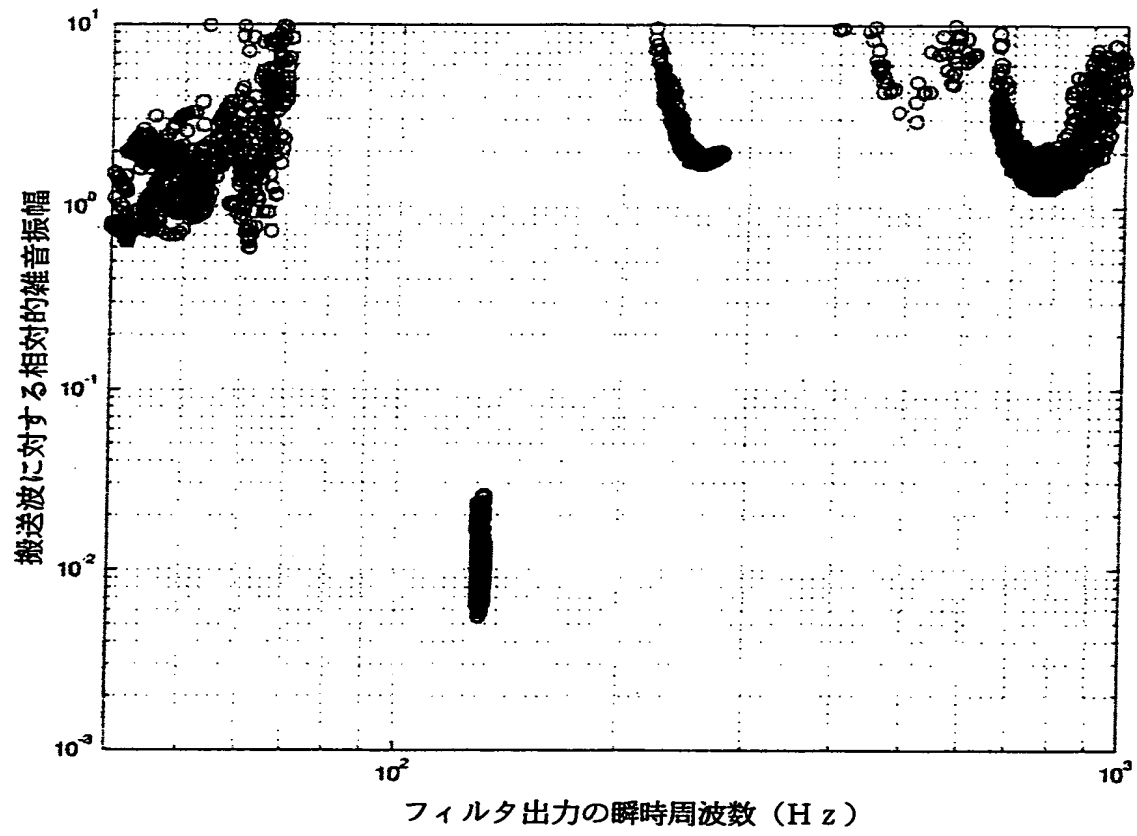


【図 8】

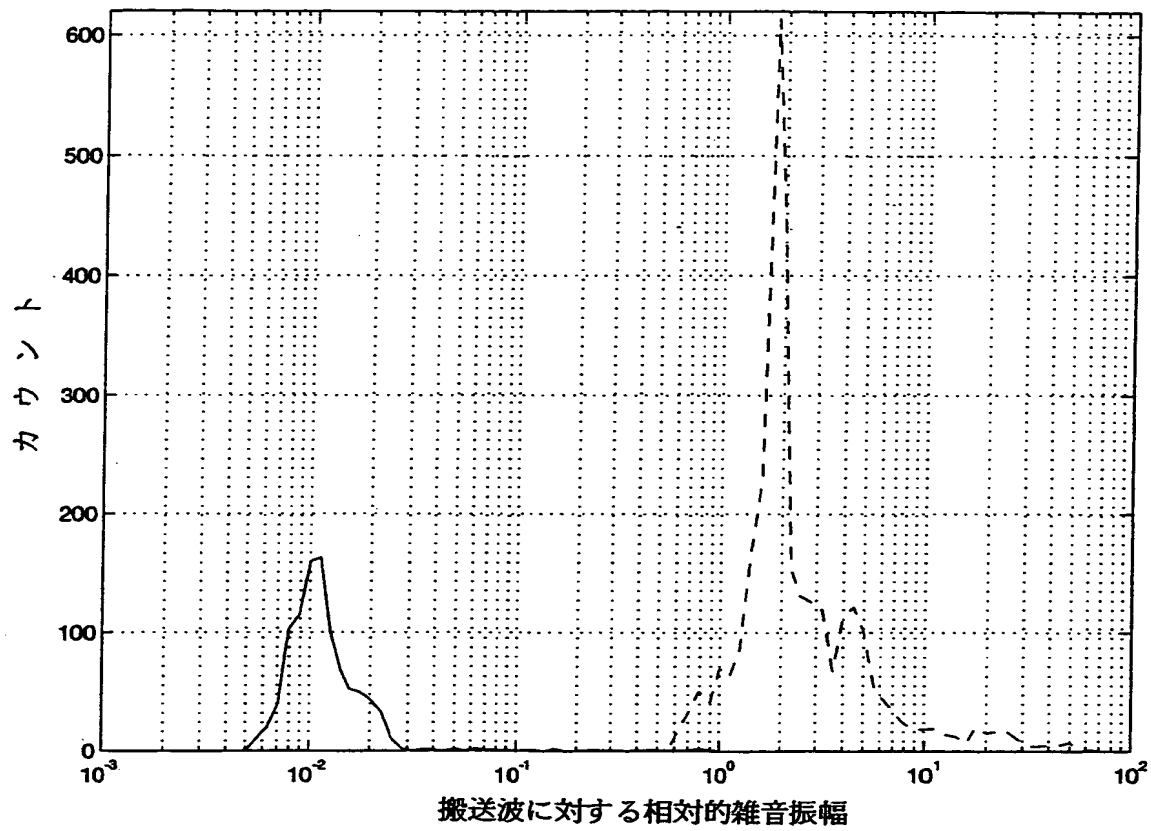




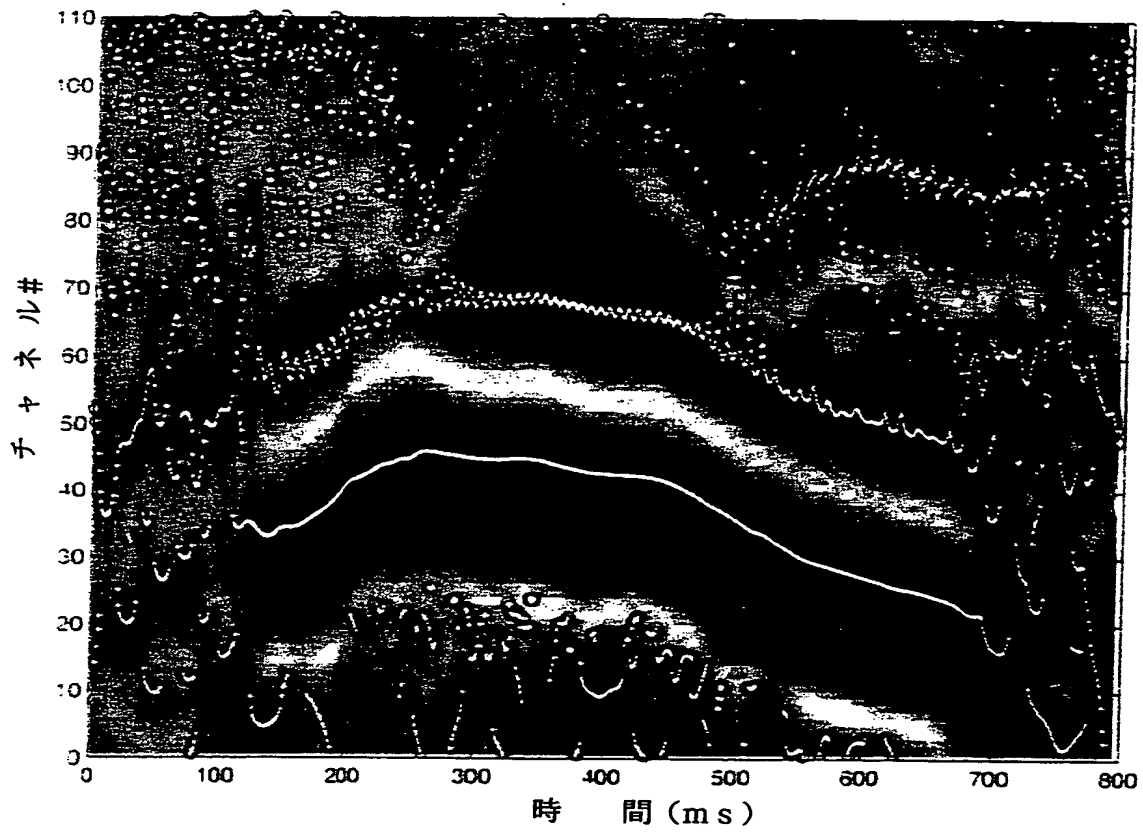
【図 9】



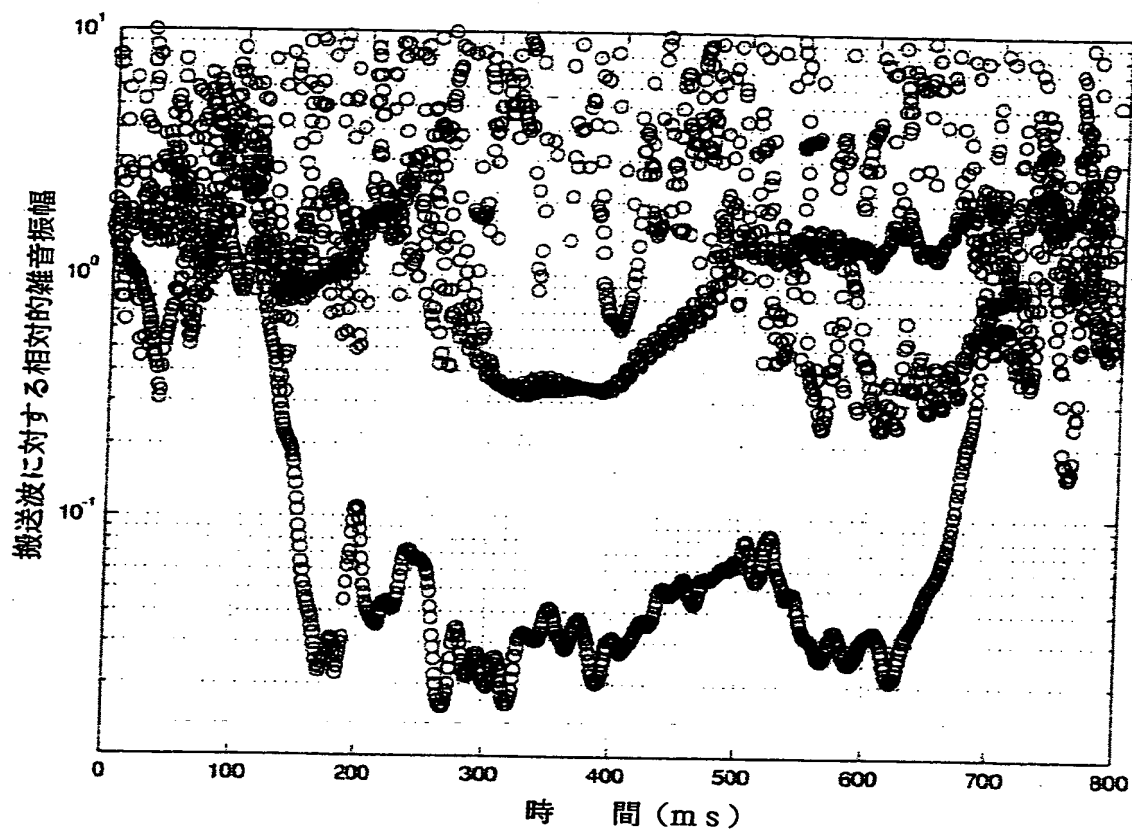
【図 10】



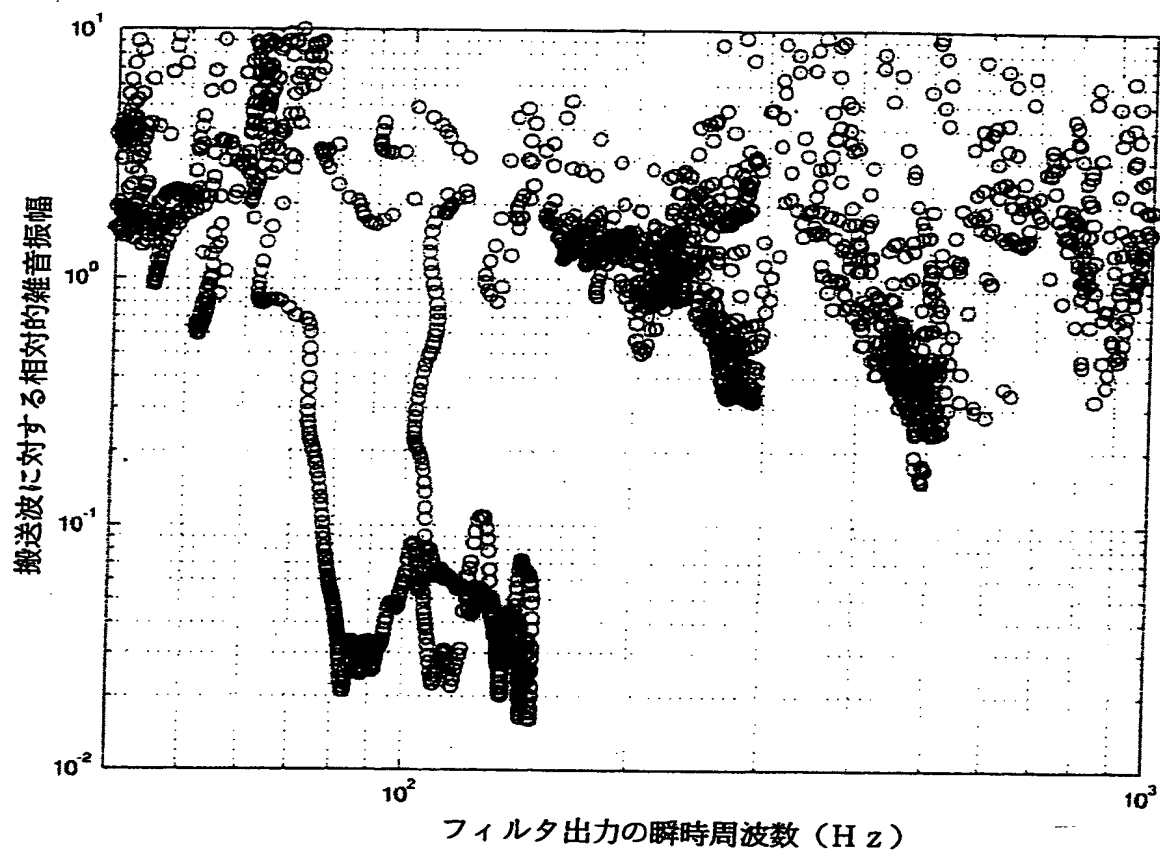
【図 11】



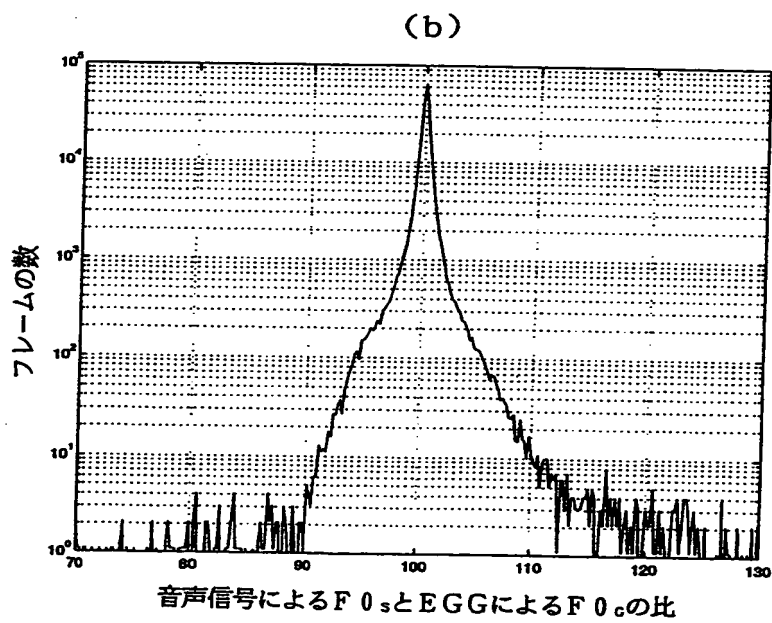
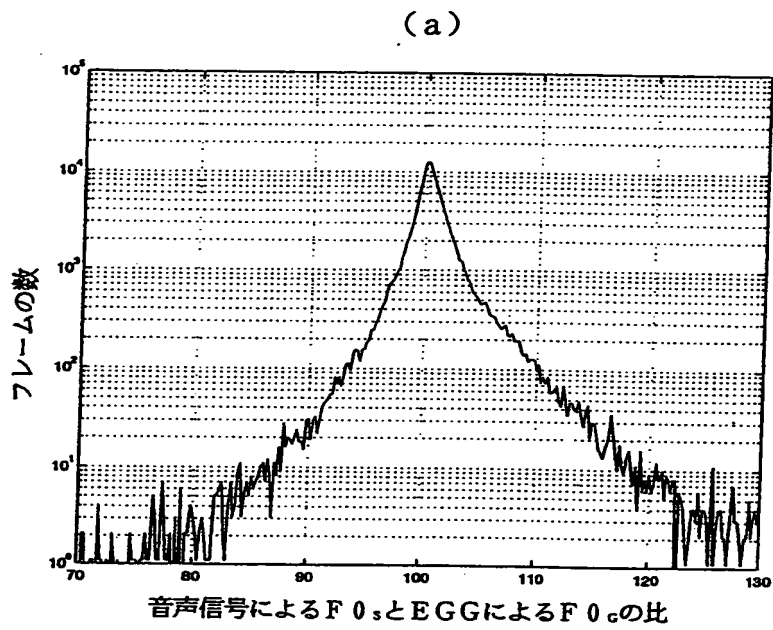
【図 12】



【図 13】



【図 14】



【書類名】

要約書

【要約】

【課題】 フィルタ中心周波数から出力の瞬時周波数への不動点の性質を瞬時のデータから定量的に解釈の明瞭な量として検出することができる音源情報の抽出方法を提供する。

【解決手段】 周波数から瞬時周波数への写像の不動点を用いた音源情報の抽出方法において、

各フィルタ 2, 9 について、瞬時周波数周波数微分回路 3, 10 により得られる瞬時周波数の周波数方向の偏微分と、瞬時周波数時間周波数微分回路 4, 11 により得られる各フィルタ出力の周波数方向の偏微分を時間方向に偏微分した値に適切な加重をかけて、時間方向に短時間の加重付きの積分を行うことにより、各フィルタについての搬送波対雑音比の推定値を搬送波対雑音比計算回路 5, 12 により計算し、搬送波対雑音比を求め、評価量の推定値を得る。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [396020800]

1. 変更年月日	1998年 2月24日
[変更理由]	名称変更
住 所	埼玉県川口市本町4丁目1番8号
氏 名	科学技術振興事業団



出 願 人 履 歴 情 報

識別番号 [592179296]

- |          |                        |
|----------|------------------------|
| 1. 変更年月日 | 1992年 8月19日            |
| [変更理由]   | 新規登録                   |
| 住 所      | 京都府相楽郡精華町大字乾谷小字三平谷5番地  |
| 氏 名      | 株式会社エイ・ティ・アール人間情報通信研究所 |
|          |                        |
| 2. 変更年月日 | 2000年 3月27日            |
| [変更理由]   | 住所変更                   |
| 住 所      | 京都府相楽郡精華町光台二丁目2番地2     |
| 氏 名      | 株式会社エイ・ティ・アール人間情報通信研究所 |

